MEASURES OF VARIABILITY

By E. J. G. Bradford, King's College, London

When the same test is applied several times the individual tested may vary at each application, in the efficiency of his response. The standard measures of variability,—the Average Deviation, the Standard Deviation, and the Coefficient of Variation,—can all be used to measure the degree of the individual's variability. But all these measures, however efficient they may be for measuring variability within a group of individuals, have serious disadvantages when applied to the variability of the individual.

Since all three measures suffer from the same defects, a consideration of the Average Deviation (A. D.) may serve as a basis for argument. In the first place the deviations are calculated from the mean performance. Continued application of the same psychological test enables the individual tested to improve in efficiency. Thus, for example, the mean performance at the last two applications will be higher than the mean performance at the first two. There is, in fact, a general improvement in the mean performance,—or stated differently, the mean is progressive.

A more exact measure of variability, then, would be one in which the deviations were calculated from a progressive mean instead of from a fixed mean.

The diagram below represents the scores of two individuals A and B at ten consecutive applications of a controlled association test, the exact nature of which will be described later.

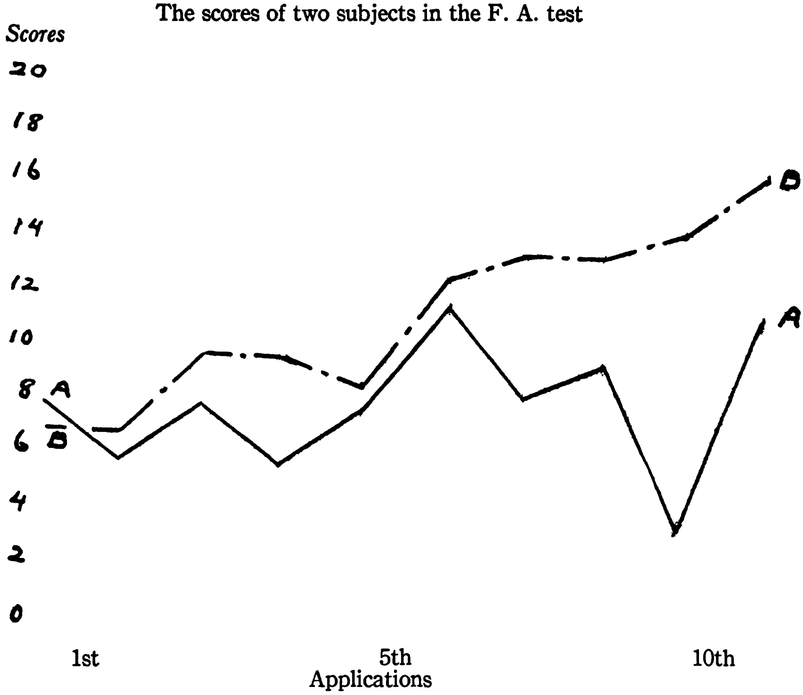
What strikes one on looking at the two series of results, is that subject A is more variable than subject B; for the line A certainly presents a far more irregular appearance than does line B, yet the means and A. D.'s of the two series are as follows:

A. Mean 8 [nearest whole number] A. D. 1.9
B. Mean 12 [nearest whole number] A. D. 3.0

Subject B shows a much greater improvement with practice than does subject A, and it is this improvement which results in such a large A. D. Herein lies the second great

DIAGRAM I

The scores of two subjects in the F. A. test



disadvantage of the A. D. If for psychological purposes we are to distinguish between variability and improvability, as seems advisable, the A. D. (or Mean Variation) must be discarded as a measure of individual variability. What is required therefore, is a coefficient of variability, which gives a measure of variability quite independently of the subject's improvability.

Consider the following case. A subject makes the following ten consecutive scores:

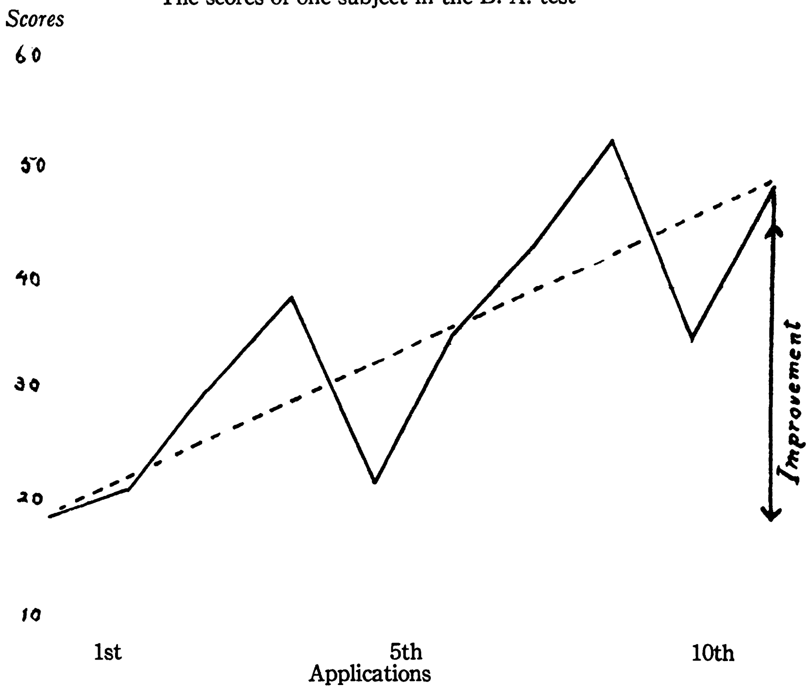
19, 21, 30, 39, 21, 35, 44, 53, 35, 49
(2) (9) (9) (18) (14) (9) (9) (18) (14)

The total variability can be measured by taking the sum of the differences (figures given in brackets above) between every two consecutive scores. In effect this amounts to reckoning the variations from a progressive mean.

It can be seen more plainly from the graphed results of these scores (Diagram II) that the sum of the differences

DIAGRAM II

The scores of one subject in the B. A. test



between the consecutive scores has been increased by the total difference between the initial and final means. In other words it has been increased by the improvement shown between the first and last measures. Hence to correct for improvability it will be necessary to diminish the total variability of the total improvement.

Let n be the number of measures.
m_1 and m_n the first and last measures respectively.
Then
$m_n - m_1$ is the improvement due to practice.
and

$$\left\{\left(m_n - m_{n-1}\right) + \left(m_{n-1} - m_{n-2}\right) + \ldots\ldots + \left(m_2 - m_1\right)\right\}$$

is the total variability due to all causes combined.

Putting
$m_n - m_1 = D$
$(m_n - m_{n-1}) = d$
We get
$(\Sigma d - D)$ as the total variability due to causes other than practice.

For a small number of measures, or at least such a number as shall be well within the limits of improvement, we can take as the measure of improvability (i)

$$i = \frac{m_n - m_1}{n} \text{ or } \frac{D}{n}$$

As a measure of variability including improvement (vi) we have

$$vi = \frac{1}{n}\left\{\Sigma\left(m_n - m_{n-1}\right)\right\} \text{ or } \frac{\Sigma d}{n}$$

Then it follows that as a measure of pure variability (v) we get

$$v = \frac{1}{n}\left\{\Sigma\left(m_n - m_{n-1}\right) - \left(m_n - m_1\right)\right\}$$

or,

$$v = \frac{1}{n}\left\{\Sigma d - D\right\}.$$

The actual scores of subjects A and B (Diagram I) were:

A. 8, 6, 8, 6, 8, 12, 9, 10, 4, 12.
B. 7, 7, 10, 10, 9, 13, 14, 14, 15, 17.

The coefficients work out as follows:

	A.	B.
Mean..................	8.0	12.0
A. D..................	1.9	3.0
vi....................	3.0	1.2
i......................	.4	1.0
v......................	2.6	.2

From these two cases it would appear that the A. D. is more a measure of improvability than of variability. This problem as to the relations between the different coefficients will be dealt with in the sequel.

To show the relations between the various coefficients of variability and incidentally their relations to ability, certain experiments were made, the methods and results of which are given below.

The Subjects. Two classes of boys from a good elementary school in London were tested. Of these classes, in scholastic attainments, Class 4 was a year in advance of Class 5. In both classes the average age was 11.0 years, for Class 4 $\sigma = .60$ years and class 5 $\sigma = .95$ years. Hence, although of the same average age, the more advanced class was the more homogenous with regard to age. The number of boys who completed the whole series of tests were 43 in Class 4 and 45 in Class 5.

The Tests. The tests employed were the Backward-alphabet test and the Forward-alphabet test. (Hereafter referred to as the B. A. and F. A. tests respectively.) The former test needs no description in that it is described by Whipple [1] and has been used by other investigators.[2] The latter test differs from the former only in that the letter succeeding the stimulus letter is required instead of the letter preceding the stimulus letter.

Procedure. The tests were applied three times a week, on Mondays, Wednesdays and Fridays, for a period extending over about a month. The subjects were supplied with a paper on which twenty letters were printed in a vertical column. There were three different orders for the letters so that there was an interval of a whole week between each application of the same list.

At the commencement of this series of tests 70 seconds was allowed for the B. A. and 35 seconds for the F. A. test. The time was gradually cut down till at the end the times were 35 and 25 seconds respectively. The scores were weighed proportionally to the time given, in each case bringing them up to the 70 and 35 second standards respectively.

The Relation Between the Two Tests. The table below gives the correlation coefficients between the B. A. and F. A. tests, as regards mean ability, variability and improvability.

TABLE 1

	Mean	A. D.	*vi*	*i*	*v*
Class 4	.77	.24	.42	.34	.40
Class 5	.86	.35	.32	.32	[*.04*]

For Class 4:
$v = 3$ P. E. when $v = .29$
$v =$ P. E. when $v = .11$
For Class 5:
$v = 3$ P. E. when $v = .28$
$v =$ P. E. when $v = .10$

In the above and subsequent tables of correlation coefficients, those coefficients which are less than three times the P. E.

[1] *Manual of Mental and Physical Tests.* 1910, 326.

[2] H. A. Aikins, E. L. Thorndike and E. Hubbell, *Psychol. Rev.* IX., 1902, 374 ff.

are printed in italics, and those which are less than the probable error are put in brackets.

It can be seen in the above table that there is a high correlation between the mean ability in the B. A. and F. A. tests. This is what might have been expected from the similar nature of the tests. Improvability (as measured by i), in the one test correlates with improvability in the other. The degree of correlation is not high but it is at least deserving of consideration since it is as great as three times the probable error; and it is approximately the same for both groups of subjects.

The coefficients obtained by correlating the variability (v) for the two tests shows that in Class 4 there is a marked tendency for variability in the B. A. test to be accompanied by variability in the F. A. test. On the other hand there is a marked absence of a similar correlation for Class 5.

Since both the i values and the v values show an appreciable degree of correlation between the two tests for Class 4, this would be sufficient to account for the fact that the vi values give a higher correlation than either of the other two separately. For Class 5 the vi values do not give greater correlation, this may be due to the fact that the v values for this class show practically no correlation at all.

The Relation between the Means and A. D.'s.—Table 2 shows that a high correlation obtains between the A. D. and the Mean in both tests and for both classes.

TABLE 2

	Correlation between Mean and A. D.	
	B. A.	F. A.
Class 4	.63	.50
Class 5	.40	.59

Does this high correlation mean that, generally speaking, variability is a function of ability; that the more able an individual is, the more variable he will be?

An answer to this question will be found in Table 3, which gives the correlation coefficients between the other measures of variability and improvability.

TABLE 3

	B. A. test			F. A. test		
	Correlation between Mean and:—					
	vi	*i*	*v*	*vi*	*i*	*v*
Class 4..............	[.*03*]	.48	—.*24*	[.*07*]	.38	—.*11*
Class 5..............	.*23*	.33	[.*04*]	.43	.74	[.*00*]

It is obvious that the answer to the above question is in the negative. For these two tests and for these two groups of individuals there is no marked tendency for variability to correlate with ability, either positively or negatively. None of the coefficients are as great as three times the probable error. The apparent exception to this generalization is found in the correlation between the Mean and *vi* for Class 5 in the F. A. test. This positive correlation is probably accounted for by the fact that the improvability factor which the *vi* coefficient contains, itself correlates very highly with the Mean.

The second point of note in this table is that in all four cases ability correlates with improvability; all four coefficients are greater than three times the P. E.

The Relation of Variability and Improvability with the A. D. A question which arises from the above considerations is, why should the A. D. correlate highly with the Mean, while *vi* and *v*, the other measures of variability do not?

TABLE 4

	B. A.			F. A.		
	Correlations between A. D. and:—					
	vi	*i*	*v*	*vi*	*i*	*v*
Class 4..............	.*16*	.71	—.31	.47	.66	.*12*
Class 5..............	.58	.61	.*24*	.52	.59	.*17*

The high correlations between A. D. and *i* and the low correlations between A. D. and *v*, show that the A. D. is not a measure of variability in the limited sense in which the term here is used. The A. D. is far more a measure of improvability than of variability.

A complete table of the correlation coefficients between the various measures and the Mean is given below.

TABLE 5

	Coefficients	B. A.		F. A.	
		Mean	A. D.	Mean	A. D.
Class 4	A. D.	.63	...	.50	...
	i	.48	.71	.38	.66
	vi	[*.03*]	*.16*	[*.07*]	.47
	v	*—.24*	—.31	*—.11*	***.12***
Class 5	A.D.	.40	...	.59	...
	i	.33	.61	.74	.59
	vi	*.23*	.58	.43	.52
	v	[*.04*]	*.24*	[*.00*]	***.17***

The Relation Between Variability and Improvability. The fact that the *vi* value correlates less highly with the mean value than does the *i* value, suggests that the variability factor in the *vi* coefficient acts in opposition to the improvability factor. This view is well substantiated by the correlation coefficients between the *i* values and the *v* values. Generally speaking, it would appear that the greater a subject's improvability the less would be his variability.

TABLE 6

	Correlations between *i* and *v*	
	B. A.	F. A.
Class 4	—.35	—.61
Class 5	*—.24*	—.28

Conclusions

1. The A. D. of a series of performances by the same individual, is not a measure of variability but rather one of improvability.

2. A relation of a positive nature exists between ability and improvability.

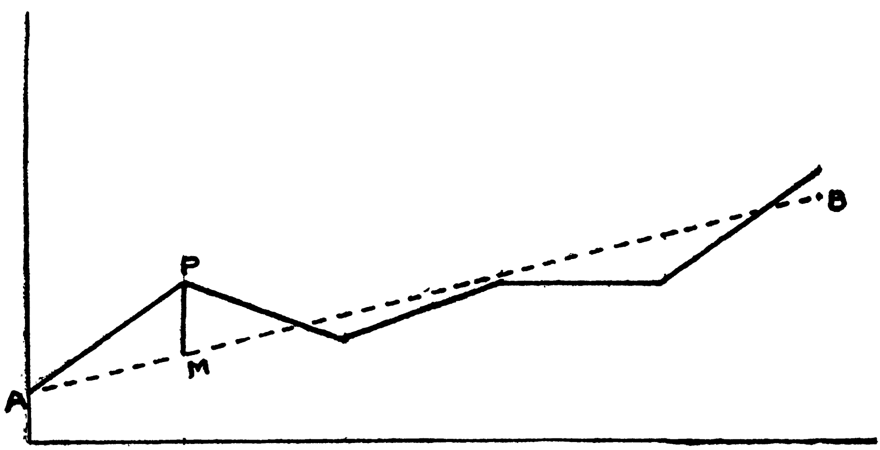
3. No such relation exists between ability and variability (v values) of performance in the two tests used.

4. The relation between variability and improvability is of a negative nature; that is, so far as these two groups of subjects and these two tests are concerned.

Appendix

An objection to the measures of variability suggested in this paper, is that they do not fall into line with those used in statistical work. To meet this objection Dr. W. Brown has made out a formula "to determine the true value of the

Diagram III



M. V. of a series of performances of a single individual, when the performances show a progressive improvement."

Let AB be the best fitting straight line, its equation being $g = mx + c$. Then the true M. V. will be the average of all such distances as P. M., taken without regard to sign.

Put PN = Y, MN = y = mx + C

Using the method of best squares we get as the best fitting line,

$$y = \frac{12S\,(Yx) - 6\,(x+1)\,S\,(Y)}{x(x^2-1)}.x_x + \frac{(4x+2)\,S\,(Y) - 6S\,(Yx)}{x(x-1)}$$

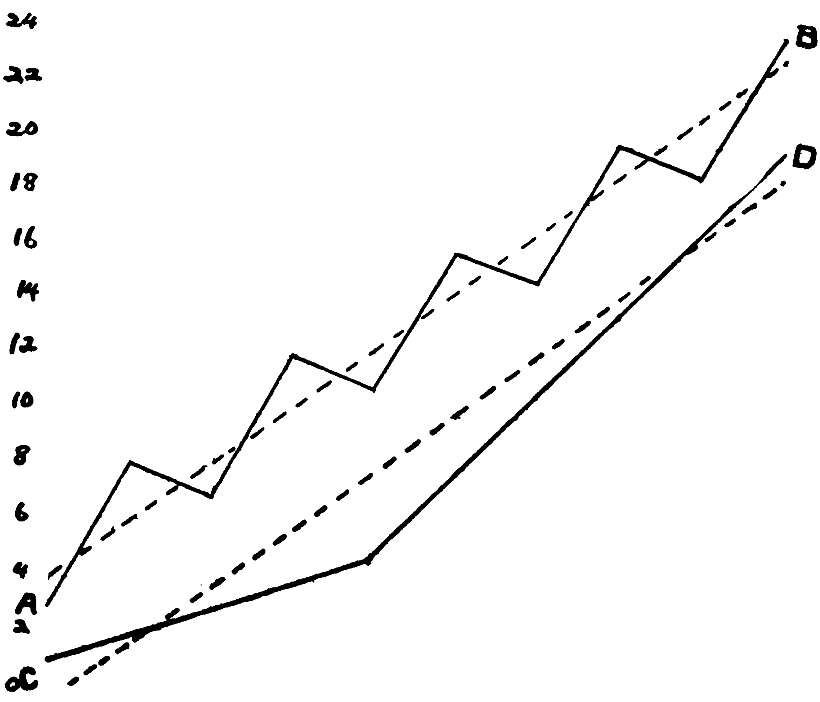
where x_x is the current coordinate.

The values of the successive ordinates of this line are determined by substitution; then the average of the differences between these ordinates and the corresponding ones of the curve (taken as all positive) will be the true value of the M. V.

Dr. Brown's "true M. V." is however a measure of the change of *the rate of improvement,* as well as of variability between consecutive performances (considered apart from improvement).

The two imaginary cases given below show this difference. The Curve CD shows change in the rate of improvement. The curve AB shows variability between consecutive performances.

DIAGRAM IV



The performances represented by the line AB give a "true M. V." of 1.3 and v=.8. From the curve CD the "true M. V." works out to 1.3 and v = O.

In the case of the line AB the "true M. V." measure the variability between consecutive performances only. The "true M. V." of the curve CD is a measure of the change in the rate of improvement.

The "true M. V." of each of the 45 subjects for the B. A. test (Class 5) was calculated and the results were correlated with the other measures with the following results:

CORRELATION COEFFICIENTS BETWEEN "TRUE M. V." AND:—

v	*i*	*vi*	Mean	A. D.
.48	.37	.74	.34	.58